DOI:10.11931/guihaia.gxzw201907001

油棕 WRKY 转录因子的全基因组鉴定与分析

周丽霞，曹红星*

（中国热带农业科学院椰子研究所/海南省热带油料作物生物学重点实验室，海南　文昌　571339）

摘　要：该研究从 NCBI 网站下载油棕全基因组序列信息，从 Arabidopsis Information Resource（TAIR）数据库中下载得到拟南芥 WRKY 转录因子序列，并在油棕基因组数据库中进行 BLAST 同源序列比对分析，通过 NCBI 在线工具 CDD 和 PFAM 数据库进行蛋白结构与分析，剔除无 WRKY 结构域的系列，利用生物信息学方法对油棕 WRKY 转录因子进行分析及功能预测。结果表明：从油棕基因组数据库中发掘 WRKY 转录因子 95 个，该 WRKY 转录因子蛋白质所编码氨基酸大小为 116～1 303 bp，95 个均为亲水性蛋白，总体为不稳定蛋白（EgWRKY25 和 EgWRKY56 除外），60 个蛋白以 α-螺旋为主要二级结构元件，35 个以无规卷曲为主要二级结构元件。保守结构域系统进化树结果表明，油棕 WRKY 转录因子家族蛋白主要分为三大类（I、II 和 III），I 类分为 IC 和 IN 亚类，II 类分为 II a、II b、II c 和 II d 亚类。内含子和外显子结构结果表明，EgWRKY 基因结构进化高度保守。该研究结果对油棕 WRKY 转录因子的挖掘、功能分析及分子生物学研究奠定了基础，此外，还对分子育种和遗传改良提供参考。

关键字：油棕，基因组，WRKY 转录因子，生物信息学，表达分析

Genome-wide identification and analysis for oil palm WRKY transcription factors

ZHOU Lixia, CAO Hongxing*

基金项目：农业农村部物种品种资源保护费项目（125163015000160004）；热带木本油料产业技术创新团队项目（17CXTD-13）；农业农村部农业国际交流与合作项目（SYZ2019-12）；国家自然科学基金（31601723）[Supported by Protection Fee of Species Resources in Agriculture and Rural Area（125163015000160004）; Tropical Woody Oil Industry Technology Innovation Team Project（17CXTD-13）; Agricultural International Exchange and Cooperation Project of the Ministry of Agriculture and Rural Areas（SYZ2019-12）; the National Natural Science Foundation of China （31601723）]。

作者简介：周丽霞（1982-），女，山东聊城人，硕士研究生，研究方向为热带棕榈作物分子生物学，（Email）glzz_2009@163.com。

***通信作者：**曹红星，博士，研究员，研究方向为热带油料作物育种及栽培研究，（E-mail）411680755 @qq.com。

(*Coconut Research Institute*, *Chinese Academy of Tropical Agricultural Science/Hainan Provincial Key Laboratory of Tropical Oil Crops Biology*, Wenchang 571339, Hainan, China)

Abstract: This study downloaded the whole genome sequence information of oil palm from NCBI website, downloaded the WRKY transcription factor sequence from Arabidopsis Information Resource (TAIR) database, and carried out BLAST homology sequence alignment analysis in oil palm genome database. Protein structure and analysis were carried out through CDD and PFAM database of NCBI online tool, and the series without WRKY domain were eliminated. Bioinformatics Analysis and Functional Prediction of WRKY Transcription Factor in Oil Palm. The results showed that 95 EgWRKY transcription factors were excavated, and they encoded 116−1 303 bp amino acids, predicting hydrophilic and unstable (except for EgWRKY25 and EgWRKY56). The main structure of 60 EgWRKY proteins was α-helix, and the remaining 35 proteins were irregular curl. The phylogenetic analysis on conserved domain showed that EgWRKY transcription factor family proteins were divided into three categories (I, II and III). Category I could be separated into IC and IN, and category II was classified into II a, II b, II c and II d. The intron-exon structure analysis revealed that structures of EgWRKY gene were highly conserved. This research will lay a foundation for the study of molecularbiology of oil palm and provide references for its genetic modification and molecular breeding.

Key words: Oil palm, genome, WRKY transcription factors, bioinformatics, gene expression

植物经过长期的进化，自身形成了一套对抗不良环境的适应机制，植物受到胁迫后，通过相应的信号传递途径，诱导相关基因的表达以抵御胁迫，而基因的表达受到转录因子的调控（谢政文等，2016）。转录因子又称反式作用因子，其作为一类重要的调控基因，主要通过与基因启动子区域中的顺式作用元件结合来发挥调控作用，近年来已成为基因挖掘与功能分析领域的研究热点（贾翠玲和侯和胜，2010）。转录因子的特点为含有 DNA 结合域、转录调控区、核定位信号区及寡聚化位点等结构。根据 DNA 结合域结构的不同，转录因子又可分为 MYB、WRKY、NAC、Zinx finger（锌指蛋白）等诸多家族（伍林涛等，2013）。

WRKY 作为植物中最大的转录因子家族之一，通过结合植物次生代谢产物合成途径关键酶基因的启动子元件来调控植物的代谢过程（丁蒙蒙等，2018）。WRKY 的 N-末端具有 7 个高度保守氨基酸残基 WRKYGQK，C 端有 1 个非典型的锌指结构，其结构域由近 60 个氨基酸组成。根据 WRKY 结构域的数目及锌指结构的类型，WRKY 蛋白可分为 3 个类型，第一类为含有 2 个 WRKY 结构域，锌指结构的类型为 C_2H_2，第二类为含有 1 个 WRKY 结构域，锌指结构的类型为 C_2H_2，第三类为含有 1 个 WRKY 结构域，锌指结构的类型为 C_2HC（伍林涛等，2013）。WRKY cDNA 最初是从甘薯中克隆出来的（Ishiguro & Nakamura, 1994），随后在其他植物如欧芹（Rushton et al., 1995）、野燕麦（Rushton et al., 1996）、拟南芥（Pater et al.,1996）、水稻（孙

利军等，2014）、枣树（Xue et al., 2019）等不同物种中分离鉴定出来。谷彦冰等（2016）利用 WRKY 保守域全蛋白序列鉴定出 61 个桃 WRKY 基因，通过生物信息学分析发现桃 WRKY 蛋白分为 I、II 和 III 类型，应用半定量和荧光定量 PCR 技术发现有 16 个 WRKY 基因均在桃的根、茎、叶、花和果重表达。刘潮等（2017）基于桑树全基因组蛋白数据库，鉴定出 55 个桑树 WRKY 基因，通过系统进化分析将 WRKY 蛋白分为 I、II 和 III 类型，应用保守结构域分析发现 WRKY 蛋白序列高度保守，在植物抵御非生物胁迫过程中发挥作用。包昌艳等（2018）应用"红阳"猕猴桃全基因组数据，鉴定出 89 个 WRKY 基因，通过进化分析发现 WRKY 蛋白可分为 I、II 和 III 类型，有 33 个 WRKY 基因在猕猴桃根、叶、花和果四个器官中均有显著表达。Fei et al.（2019）从陇南大红袍花椒中分离鉴定出 38 个 WRKY 家族成员，其中 *ZbWRKY33* 是对干旱胁迫反应最敏感的成员之一，通过直接与乙烯合成前体基因 asc 结合调节花椒抵御干旱的能力。由此可见，诸多植物中 WRKY 转录因子均得到鉴定与分析，但目前针对油棕 WRKY 转录因子基因及蛋白质的鉴定与生物信息学分析的研究鲜见报道，本研究基于油棕基因组数据，利用生物信息学方法全面分析油棕 WRKY 转录因子家族结构及特征，为深入研究 WRKY 转录因子家族的生物学功能奠定基础。

1 材料与方法

1.1 EgWRKY 序列获取与鉴定

油棕全基因组序列从 NCBI 下载获得，拟南芥 WRKY 基因序列从 Arabidopsis Information Resource（TAIR）数据库中下载获得（http://www.arabidopsis.org），以拟南芥 WRKY 序列为探针，阈值设定为 e^{-10}，在油棕基因组数据库中进行同源序列比对分析得到油棕 WRKY 基因序列，将得到的油棕 WRKY 基因序列在美国麻省理工学院在线基因扫描服务器（The GENSCAN Web Server at MIT）搜索以获得 WRKY 蛋白序列，通过 NCBI 在线工具 CDD（http://www.ncbi.nlm.nih.gov/cdd）和 PFAM 数据库（http://pfam.xfam.org/）进行蛋白结构域分析，剔除无 WRKY 保守结构域的蛋白序列。

1.2 方法

1.2.1 EgWRKY 蛋白理化性质分析

利用 ProtParam 分析 EgWRKY 蛋白的理化性质，利用 NetPhos 3.1 Server 和 DictyOGlyc 1.1 Server 预测氨基酸序列磷酸化及 O-糖基化修饰情况，利用 ProtScale 分析蛋白质的亲/疏水性，利用 SignalP 和 ProtComp 预测蛋白的信号肽和亚细胞定位预测，应用 SOPMA 预测蛋白的二级结构。

1.2.2 EgWRKY 的进化树、外显子及内含子分析

应用 MEGA 6.06 软件构建 EgWRKY 进化树，执行参数为 Neighbor-Joining，Bootstrap 重复 1 000 次。

2 结果与分析

2.1 EgWRKY 转录因子家族成员基本信息

基于油棕全基因组数据及拟南芥 WRKY 基因序列，通过 blast 搜索同源序列和 WRKY 蛋白保守结构域鉴定，共挖掘 95 个油棕 WRKY 转录因子成员，该 95 个 WRKY 转录因子的基因序列号及 cDNA 序列信息见附件 1；由表 1 可知，该 95 个蛋白质所含氨基酸大小为 116～1 303 bp，其中蛋白氨基酸数目小于 300 aa 的基因序列占 35.8%，介于 300～700 aa 的基因序列占 57.9%，大于 700 aa 的基因序列占 6.3%。

表 1 油棕 WRKY 转录因子家族成员基本信息

Table 1 Information of oil palm WRKY transcription factor family genes

基因名称 Gene name	染色体定位 Chromosomal localization	蛋白登录号 Protein accession No.	氨基酸大小 Peptide length（aa）
EgWRKY01	scaffold_1:2828118.. 2834829 -	XP_010927976.1	237
EgWRKY02	scaffold_1:3528352..3531492 +	XP_010934629.1	517
EgWRKY03	scaffold_1:3581086..3582588 -	XP_010935352.1	196
EgWRKY04	scaffold_1:3587901..3589580 +	XP_010935437.1	206
EgWRKY05	scaffold_1:3597496..3599292 -	XP_010912931.1	202
EgWRKY06	scaffold_1:3790630..3791759 -	XP_010936886.1	176
EgWRKY07	scaffold_1:8508520..8510767 -	XP_010913954.1	356
EgWRKY08	scaffold_1:10329250..10329726 -	XP_010917440.1	158
EgWRKY09	scaffold_1:10334215..10335866 -	XP_010916830.1	346
EgWRKY10	scaffold_1:10559780..10561375 +	XP_010916734.1	276
EgWRKY11	scaffold_1:18267529..18269693 +	XP_010926916.1	342
EgWRKY12	scaffold_1:19074456..19077025 +	XP_010926185.1	588
EgWRKY13	scaffold_1:45150214..45153149 -	XP_010939557.1	258
EgWRKY14	scaffold_1:46090220..46106758 +	XP_010939487.1	947

EgWRKY15	scaffold_2:1244849..1249774 +	XP_010909377.1	500
EgWRKY16	scaffold_2: 18378213..18380205 -	XP_010912156.1	304
EgWRKY17	scaffold_2: 34420487..34423323 -	XP_010913024.1	181
EgWRKY18	scaffold_2:38960111..38968368 -	XP_010913249.1	738
EgWRKY19	scaffold_2:47764368..47766219 +	XP_010931316.1	316
EgWRKY20	scaffold_2:49916408..49921049 -	XP_010914024.1	739
EgWRKY21	scaffold_2:60595545..60597170 +	XP_010914568.1	254
EgWRKY22	scaffold_3:2897890..2903252 -	XP_010915067.1	240
EgWRKY23	scaffold_3:3976191..3981682 +	XP_010915156.1	632
EgWRKY24	scaffold_3:4421975..4423228 -	XP_010915168.1	208
EgWRKY25	scaffold_3: 4455296..4457347 +	XP_010915169.1	171
EgWRKY26	scaffold_3:11542191..11545474 +	XP_010915678.1	373
EgWRKY27	scaffold_3:12905634..12907531 -	XP_010915790.1	374
EgWRKY28	scaffold_3:16840840..16843289 +	XP_010916103.1	532
EgWRKY29	scaffold_3:43201571..43208037 +	XP_010917544.1	186
EgWRKY30	scaffold_3:46869949..46873630 +	XP_010917578.1	358
EgWRKY31	scaffold_4:46688084..46692604 +	XP_010919759.1	213
EgWRKY32	scaffold_4:50354451..50358395 +	XP_010920047.1	419
EgWRKY33	scaffold_5:1003267..1006117 +	XP_010920544.1	326
EgWRKY34	scaffold_5:5971453..5974564 -	XP_010921056.1	594
EgWRKY35	scaffold_5:7468889..7471135 -	XP_010921156.1	291
EgWRKY36	scaffold_5:7509255..7511846 -	XP_010921157.1	276
EgWRKY37	scaffold_5:7516052..7519209 -	XP_010921553.1	210
EgWRKY38	scaffold_5:7533908..7536374 -	XP_010921554.1	365
EgWRKY39	scaffold_5:11466843..11468463 -	XP_010921434.1	311
EgWRKY40	scaffold_5:42769327..42771373 +	XP_010922758.1	310

EgWRKY41	scaffold_5:42856852..42860147 +	XP_010922757.1	323
EgWRKY42	scaffold_5:44348667..44361739 +	XP_010923015.1	1 303
EgWRKY43	scaffold_5:47101333..47106104 +	XP_010922486.1	469
EgWRKY44	scaffold_5:48908668..48911631 -	XP_010922292.1	314
EgWRKY45	scaffold_6:3040168..3043302 -	XP_010923468.1	356
EgWRKY46	scaffold_6:4288074..4293918 -	XP_010923565.1	462
EgWRKY47	scaffold_6:4774207..4779820 +	XP_010923591.1	302
EgWRKY48	scaffold_6: 33625556..33627842 -	XP_010924340.1	575
EgWRKY49	scaffold_6:34241286..34243364 -	XP_010924354.1	336
EgWRKY50	scaffold_6: 38804476..38806170 -	XP_010924726.1	277
EgWRKY51	scaffold_6:39157107..39158687 +	XP_010924759.1	328
EgWRKY52	scaffold_6:40484853..40487251 -	XP_010924841.1	360
EgWRKY53	scaffold_6:42604078..42605690 -	XP_010925052.1	202
EgWRKY54	scaffold_6:43055219..43060457 +	XP_010925080.1	235
EgWRKY55	scaffold_7:2603710..2610166 +	XP_010925379.1	626
EgWRKY56	scaffold_7:2664908..2666305 -	XP_010925385.1	210
EgWRKY57	scaffold_7:2698084..2699488 +	XP_010925386.1	166
EgWRKY58	scaffold_7:5565923..5567535 -	XP_010925600.1	287
EgWRKY59	scaffold_7:7728041..7731472 +	XP_010925763.1	369
EgWRKY60	scaffold_7:8642979..8644669 -	XP_010925829.1	371
EgWRKY61	scaffold_7:10622308..10624799 +	XP_010926194.1	584
EgWRKY62	scaffold_7:26175565..26204615 +	XP_010926843.1	170
EgWRKY63	scaffold_7:40437708..40439232 +	XP_010927275.1	324
EgWRKY64	scaffold_8:23543798..23554103 -	XP_010928175.1	742
EgWRKY65	scaffold_8:31318423..31319807 +	XP_010928698.1	308
EgWRKY66	scaffold_8:36521373..36525318 -	XP_010929154.1	257

EgWRKY67	scaffold_8:36538138..36539837 -	XP_010929156.1	315
EgWRKY68	scaffold_8:36672320..36680988 +	XP_010929162.1	533
EgWRKY69	scaffold_8:38477347..38479734 +	XP_010929316.1	321
EgWRKY70	scaffold_9:33121964..33123639 +	XP_010930927.1	292
EgWRKY71	scaffold_10:2919960..3009126 -	XP_010931246.1	514
EgWRKY72	scaffold_10:5488658..5491941 -	XP_010931316.1	530
EgWRKY73	scaffold_10:7533000..7535705 -	XP_010931380.1	312
EgWRKY74	scaffold_10:7555119..7557529 -	XP_010931510.1	243
EgWRKY75	scaffold_10:16498194..16500066 +	XP_010931747.1	283
EgWRKY76	scaffold_10:27936570..27938051 -	XP_010932751.1	116
EgWRKY77	scaffold_10:30816717..30820483 +	XP_010932879.1	353
EgWRKY78	scaffold_11:25944218..25948814 +	XP_010933393.1	421
EgWRKY79	scaffold_14:1936977..1939615 -	XP_010937592.1	336
EgWRKY80	scaffold_14:3999586..4002656 -	XP_010937966.1	596
EgWRKY81	scaffold_14:4789612..4792391 -	XP_010938278.1	385
EgWRKY82	scaffold_14:6594036..6597123 -	XP_010938223.1	315
EgWRKY83	scaffold_15:7368664..7381113 -	XP_010939060.1	551
EgWRKY84	498..2001 +	XP_010911418.1	202
EgWRKY85	707..1876 -	XP_010911458.1	232
EgWRKY86	30951..33207 -	XP_010910323.1	133
EgWRKY87	1118523..1122926 +	XP_010904958.1	251
EgWRKY88	1154550..1156525 +	XP_010942101.1	316
EgWRKY89	776067.. 784882 -	XP_010942092.1	535
EgWRKY90	1132053.. 1196544 -	XP_010943606.1	709
EgWRKY91	1056488.. 1082285 -	XP_010905300.1	672
EgWRKY92	298091..311193 +	XP_010906869.1	551

EgWRKY93	157052..174713 +	XP_010908244.1	545
EgWRKY94	231238..253849 -	XP_010908581.1	641
EgWRKY95	245494.. 250181 +	XP_010908962.1	478

2.2 EgWRKY 蛋白理化性质分析

利用 ProtParam 在线软件（http://expasy.org/tools/protparam.html）分析 EgWRKY 蛋白的基本性质，预测结果如附件 2 所示，95 个 WRKY 蛋白中，56 个蛋白的理论等电点（pI）小于 7，39 个蛋白的 pI 大于 7，pI 最小值为 4.74（EgWRKY73），最大值为 10.18（EgWRKY84），平均 pI 为 7.15；EgWRKY25 和 EgWRKY56 蛋白质的不稳定系数小于 40，为稳定蛋白，其余 93 个 EgWRKY 蛋白的不稳定系数大于 40，为不稳定蛋白，该结果表明 EgWRKY 转录因子家族大多为不稳定蛋白；利用 NetPhos 3.1 Server 在线软件（http://www.cbs.dtu.dk/services/NetPhos/）预测氨基酸序列磷酸化，DictyOGlyc 1.1 Server 在线软件（http://www.cbs.dtu.dk/services/DictyOGlyc/）预测氨基酸序列的 O-糖基化修饰情况，发现平均磷酸化位点有 37.8 个，O-糖基化位点有 1.6 个；利用 ProtScale 在线软件（https://web.expasy.org/protscale/）分析蛋白质的亲/疏水性，从 EgWRKY 蛋白的脂肪系数来看，该 95 个 EgWRKY 蛋白的脂肪系数均小于 100，该结果表明其均为亲水性蛋白；利用 SignalP 5.0 在线软件（http://www.cbs.dtu.dk/services/SignalP/）预测蛋白的信号肽，发现 5 个 EgWRKY 蛋白存在信号肽（EgWRKY29、EgWRKY38、EgWRKY45、EgWRKY63 及 EgWRKY87），为分泌蛋白；利用 ProtComp 9.0（http://linux1.softberry.com/berry.phtml?topic= protcomppl&group=programs&subgroup=proloc）在线预测蛋白亚细胞定位，结果表明 95 个 EgWRKY 蛋白中有 84 个定位在细胞核中，占总蛋白的 88.3%，11 个在细胞质中，推测可能参与细胞质基因的转录调控；应用 SOPMA 在线软件（https://npsa-prabi.ibcp.fr/cgi–bin/np sa_automat.pl?page=npsa_sopma.html）预测蛋白的二级结构，发现所有 EgWRKY 蛋白的二级结构均有 4 种，即为：α-螺旋、β-转角、无规卷曲及延伸链，其中，60 个 EgWRKY 蛋白以 α-螺旋为主要结构，无规卷曲为次要结构，35 个以无规卷曲为主要结构，α-螺旋为次要结构，β-转角及延伸链所占比例较少。

2.3 EgWRKY 的进化树分析

利用 MEGA 6.0 对油棕 95 个 WRKY 转录因子蛋白的保守结构域进行系统进化树分析，结果如图 1 所示，油棕 WRKY 蛋白主要分为三大类，其中，根据 WRKY 保守结构域所处的位置，将第 I 类分为 I N 和 I C 两个亚类，第 II 类含有 1 个锌指结构和 1 个 WRKY 保守结构域，根据锌指结构的不同，又将其分为 II a、II b、II c 及 II d 四个亚类，第 III 类是含有 2 个 WRKY 保守结构域。

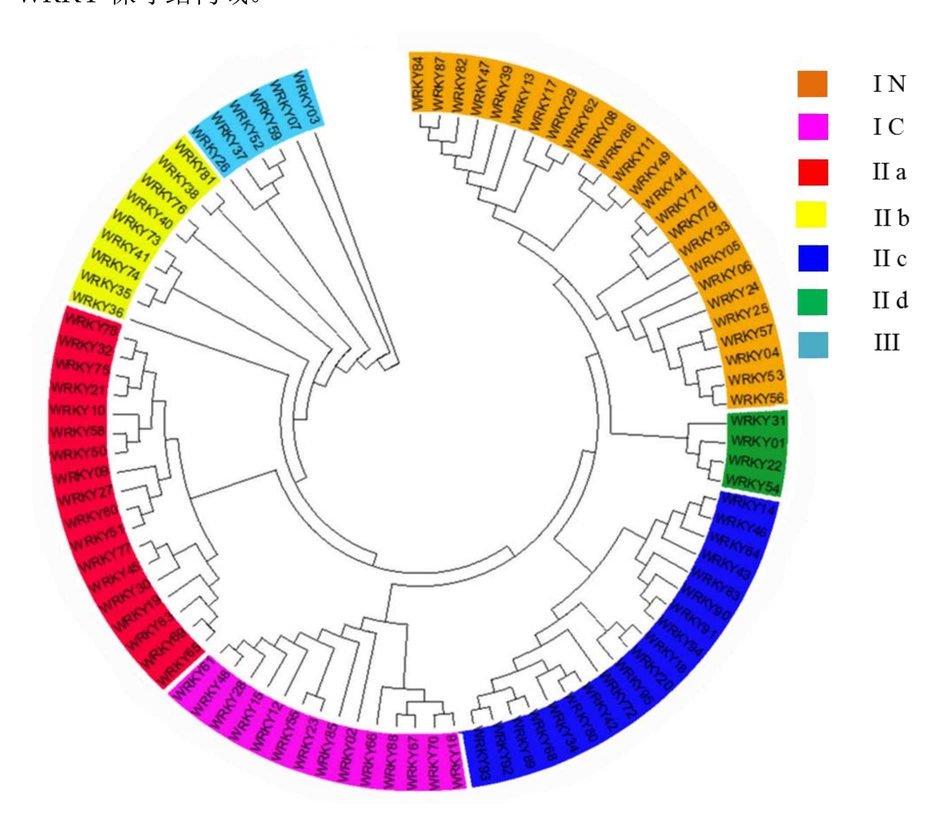
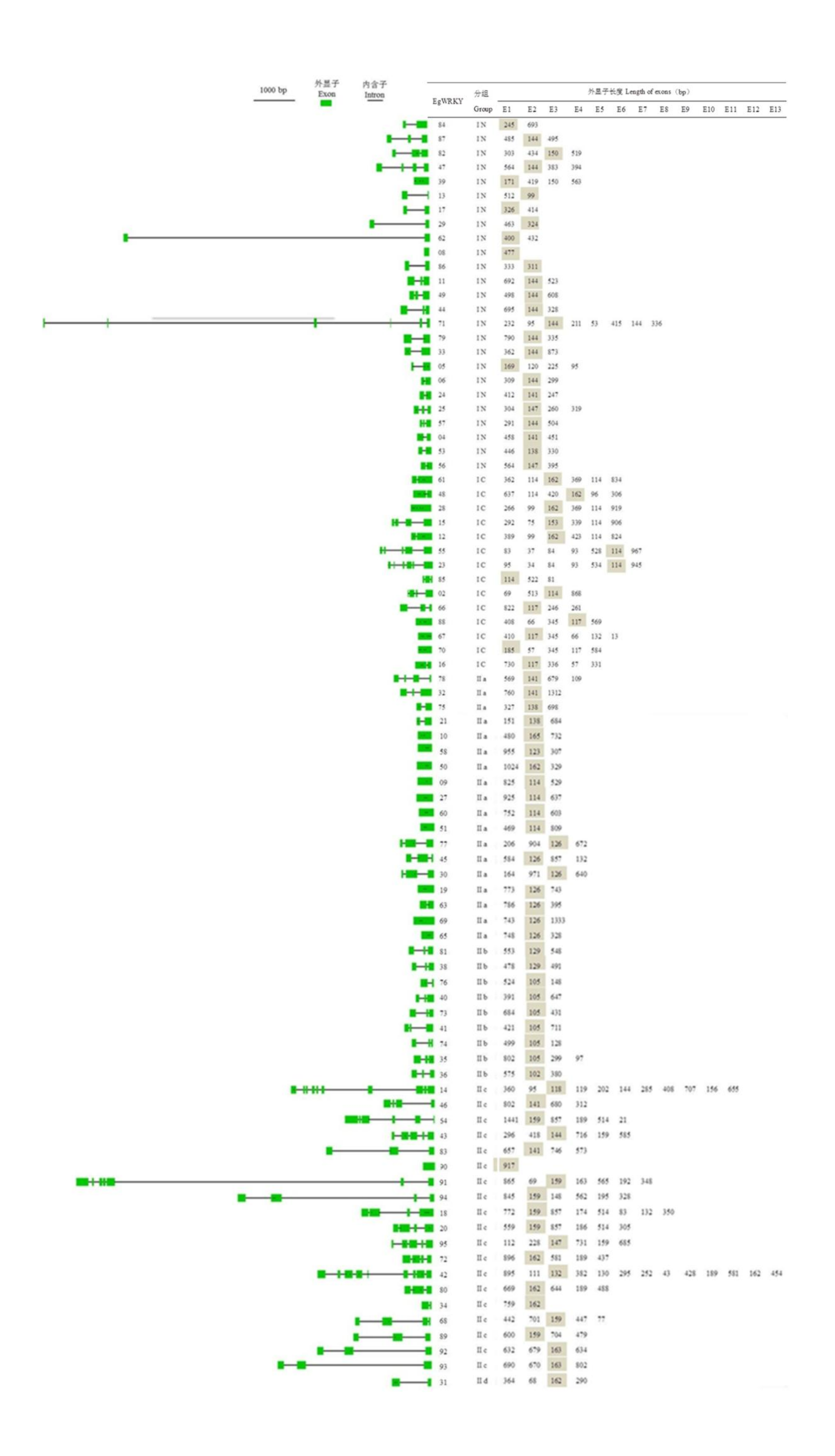


图 1 油棕 WRKY 蛋白系统进化树

Fig. 1 Phylogenetic tree of oil palm WRKY protein

2.4 油棕 WRKY 外显子及内含子

通过油棕基因组数据分析得到 WRKY 转录因子基因外显子和内含子的结构分布示意图（图 2），EgWRKY 转录因子家族成员所含内含子数的范围为 0～12 个，平均每个 EgWRKY 含内含子数为 2.99 个，其中 *EgWRKY42* 含有内含子数最多（12 个），转录因子基因长度变化范围为 477 bp（*EgWRKY08*）～89 167 bp（*EgWRKY71*），平均大小为 5 992 bp，图中阴影部分的数字代表该外显子翻译成部分 WRKY 保守结构域的长度。EgWRKY 基因家族的结构变化较大，说明在长期进化过程中油棕基因组经历了较大的变异选择。



EgWRKY	分组 Group	外显子长度 Length of exons（bp）												
		E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13
84	I N	245	693											
87	I N	485	144	495										
82	I N	303	434	150	519									
47	I N	564	144	383	394									
39	I N	171	419	150	563									
13	I N	512	99											
17	I N	326	414											
29	I N	463	324											
62	I N	400	432											
08	I N	477												
86	I N	333	311											
11	I N	692	144	523										
49	I N	498	144	608										
44	I N	695	144	328										
71	I N	232	95	144	211	53	415	144	336					
79	I N	790	144	335										
33	I N	362	144	873										
05	I N	169	120	225	95									
06	I N	309	144	299										
24	I N	412	141	247										
25	I N	304	147	260	319									
57	I N	291	144	504										
04	I N	458	141	451										
53	I N	446	138	330										
56	I N	564	147	395										
61	I C	362	114	162	369	114	834							
48	I C	637	114	420	162	96	306							
28	I C	266	99	162	369	114	919							
15	I C	292	75	153	339	114	906							
12	I C	389	99	162	423	114	824							
55	I C	83	37	84	93	528	114	967						
23	I C	95	34	84	93	534	114	945						
85	I C	114	522	81										
02	I C	69	513	114	868									
66	I C	822	117	246	261									
88	I C	408	66	345	117	569								
67	I C	410	117	345	66	132	13							
70	I C	185	57	345	117	584								
16	I C	730	117	336	57	331								
78	II a	569	141	679	109									
32	II a	760	141	1312										
75	II a	327	138	698										
21	II a	151	138	684										
10	II a	480	165	732										
58	II a	955	123	307										
50	II a	1024	162	329										
09	II a	825	114	529										
27	II a	925	114	637										
60	II a	752	114	603										
51	II a	469	114	809										
77	II a	206	904	126	672									
45	II a	584	126	857	132									
30	II a	164	971	126	640									
19	II a	773	126	743										
63	II a	786	126	395										
69	II a	743	126	1333										
65	II a	748	126	328										
81	II b	553	129	548										
38	II b	478	129	491										
76	II b	524	105	148										
40	II b	391	105	647										
73	II b	684	105	431										
41	II b	421	105	711										
74	II b	499	105	128										
35	II b	802	105	299	97									
36	II b	575	102	380										
14	II c	360	95	118	119	202	144	285	408	707	156	655		
46	II c	802	141	680	312									
54	II c	1441	159	857	189	514	21							
43	II c	296	418	144	716	159	585							
83	II c	657	141	746	573									
90	II c	917												
91	II c	865	69	159	163	565	192	348						
94	II c	845	159	148	562	195	328							
18	II c	772	159	857	174	514	83	132	350					
20	II c	559	159	857	186	514	305							
95	II c	112	228	147	731	159	685							
72	II c	896	162	581	189	437								
42	II c	895	111	132	382	130	295	252	43	428	189	581	162	454
80	II c	669	162	644	189	488								
34	II c	759	162											
68	II c	442	701	159	447	77								
89	II c	600	159	704	479									
92	II c	632	679	163	634									
93	II c	690	670	163	802									
31	II d	364	68	162	290									

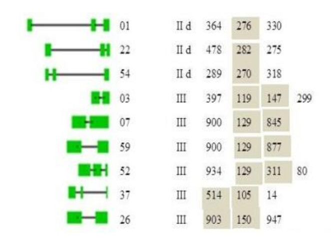


图 2 EgWRKY 基因家族外显子和内含子结构图

Fig.2 Exon-intron structures of EgWRKY gene family

3 讨论与结论

随着高通量测序技术的成熟，WRKY 作为调控植物多种生理过程的转录因子家族之一，已在多个物种中被挖掘及鉴定，如拟南芥基因组中含有 72 个 WRKY 转录因子（Ulker & Somssich, 2004），黄瓜基因组中有 132 个 WRKY 转录因子（谷彦冰等，2015），棉花基因组中有 116 个 WRKY 转录因子（Dou et al., 2014），橡胶树基因组中含有 81 个 WRKY 转录因子（Li et al., 2014），桑树基因组中含有 54 个 WRKY 转录因子（Baranwal et al., 2016）及鹰嘴豆基因组中含有 70 个 WRKY 转录因子（Waqas et al., 2019）。本研究通过分析油棕基因组数据，挖掘出 95 个 WRKY 转录因子，WRKY 家族中转录因子的数目不仅与物种的基因组相关，而且与植物在长期进化过程中所受的外界环境影响有关。油棕 WRKY 转录因子数目与其他物种相比，属于中等数目的类型，该结果可推测油棕在自然进化过程中，WRKY 基因家族受到了一定的外界环境压力。

本研究通过预测油棕 WRKY 蛋白的二级结构，发现其以 α-螺旋及无规卷曲为主，这与其他报道结果相一致（丁蒙蒙等，2018）；同时对油棕 WRKY 保守结构域进行系统进化分析，将油棕 WRKY 家族蛋白分为三大类（I、II 和 III），II 类又分为 4 个亚组。油棕 95 个 WRKY 家族成员均含有 WRKYGQK 保守基序，II 类均含有 1 个锌指结构及 1 个 WRKY 保守结构域，根据锌指结构和位置的不同又分为 4 个亚组，类似分类在拟南芥及桑树等物种中发现，该结果可推测植物 WRKY 基因家族在结构上具有高度保守性。油棕作为热带地区重要的油料作物之一，低温气候严重限制着油棕的大面积种植，若能挖掘与低温胁迫相关的调节基因，并对其功能进行预测，运用分子技术对油棕品种进行改良，提高其抗寒性，从而实现油棕大面积的种植。

参考文献：

BAO CY, DENG L, ZHOU J, et al., 2018. Genome-wide identification and analysis of the WRKY transcription factor gene family in kiwi fruit[J]. Mol Plant Breed, 2(12)：301-310. [包昌艳，邓浪，周军，等，2018. 猕猴桃 WRKY 转录因子家族全基因组鉴定与分析[J]. 分子植物育种，2（12）：301-310.]

BARANWAL VK, NEGI N, KHURANA P, 2016. Genome-wide identification and structural, functional and evolutionary analysis of WRKY components of mulberry[J]. Sci Reports, 6: 30794-30807.

DING MM, SHI XD, GU YX, et al., 2018. Transcriptome-based excavation and analysis of MYB family transcription factors in *Phoebe zhennan*[J]. Guihaia, 38(1)：90-100. [丁蒙蒙，时小东，顾雨熹，等，2018. 基于转录组的楠木 MYB 转录因子的挖掘及分析[J]. 广西植物，38（1）：90-100.]

DOU LL, ZHANG XH, PANG CY, et al., 2014. Genome-wide analysis of the WRKY gene family in cotton[J]. Mol Genet Genom, 289(6): 1103-1121.

FEI XT, HOU LX, SHI JW, et al, 2019. Patterns of drought response of 38 WRKY transcription factors of *Zanthoxylum bungeanum* Maxim[J]. Int J Mol Sci, 20(1): 68-85.

GU YB, JI ZR, CHI FM, et al., 2016. Genome-wide identification and expression analysis of the WRKY gene family in peach[J]. Hereditas, 38(3)：254-270. [谷彦冰，冀志蕊，迟福梅，等，2016. 桃 WRKY 基因家族全基因组鉴定和表达分析[J]. 遗传，38（3）：254-270.]

GU YB, JI ZR, CHI FM, et al., 2015. Bioinformatics and expression analysis of the WRKY gene family in apple[J]. Sci Agric Sin, 48(16): 3221-3238. [谷彦冰，冀志蕊，迟福梅，等，2015. 苹果 WRKY 基因家族生物信息学及表达分析[J]. 中国农业科学，48（16）：3221-3238.]

ISHIGURO S, NAKAMURA K, 1994. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5’ upstream regions of genes coding for sporamin and β-amylase from sweet potato[J]. Mol Genet Genet, 244(6)：563-571.

JIA CL, HOU HS, 2010. Structure of plant WRKY transcription factors and their roles in plant defense response[J]. Tianjin Agric Sci, 16(2)：21-26. [贾翠玲，侯和胜，2010. 植物 WRKY 转录因子的结构特点及其在植物防卫反应中的作用[J]. 植物生理与生物技术，16（2）：21-26.]

LI HL, GUO D, YANG ZP, et al., 2014. Genome-wide identification and characterization of WRKY gene family in *Hevea brasiliensis*[J]. Genomics, 104(1): 14-23.

LIU C, HAN LH, SONG PB, et al., 2017. Genome-wide identification and bioinformatics analysis for mulberry WRKY transcription factors[J]. J S Agric, 48(9)：1691-1699. [刘潮，韩利红，宋培兵，等，2017. 桑树 WRKY 转录因子的全基因组鉴定及生物信息学分析[J]. 南方农业学报，48（9）：1691-1699.]

PATER S, GRECO V, PHAM K, et al., 1996. Characterization of a zinc dependent transcriptional activator from *Arabidopsis*[J]. Nucl Acids Res，24(23)：4624-4631.

RUSHTON PJ, MACDONALD H, HUTTLY AK, et al., 1995. Members of a new family of DNA-binding proteins bind to a conserved cis-element in the promoters of α-Amy2 genes[J]. Plant Mol Biol, 29(4)：691- 702.

RUSHTON PJ, TORRES JT, PARNISKE M, et al., 1996. Interaction of elicitor- induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes[J]. EMBO J, 15 (20)：5690-5700.

SUN LJ, HUANG L, LI DY, et al., 2014. Comprehensive expression analysis suggests overlapping of rice OsWRKY transcription factor genes during abiotic stress responses[J]. Plant Physiol J, 50(11)：1651-1658. [孙利军，黄磊，李大勇，等，2014. 水稻 OsWRKY 转录因子对非生物胁迫响应的重叠表达特性分析[J]. 植物生理学报，50（11）：1651-1658.]

ULKER B, SOMSSICH IE, 2004. WRKY transcription factors: From DNA binding towards biological function[J]. Curr Opin Plant Biol, 7(5): 491-498.

WAQAS M, AZHAR MT, RANA IA, et al., 2019. Genome-wide identification and expression

analyses of WRKY transcription factor family members from chickpea (*Cicer arietinum* L.) reveal their role in abiotic stress-responses[J]. Genes Genom, 41(4): 467-481.

WU LT, DU CF, ZHANG MQ, et al., 2013. The structure and function of WRKY transcription factors in abiotic and biotic stress[J]. Mol Plant Breed, 11(4)：634-638. [伍林涛，杜才富，张敏琴，等，2013. WRKY 转录因子的结构及其在植物抗逆境胁迫中的功能[J]. 分子植物育种，11（4）：634-638.]

XIE ZW, WANG LJ, CHEN MY, et al., 2016. Studies on WRKY transcription factors and their biological functions in plants[J]. J Agric Sci Technol, 18(3)：46-54.[谢政文，王连军，陈锦洋，等，2016. 植物 WRKY 转录因子及其生物学功能研究进展[J]. 中国农业科技导报，18（3）：46-54.]

XUE C, LI H, LIU Z, et al., 2019. Genome-wide analysis of the WRKY gene family and their positive responses to phytoplasma invasion in *Chinese jujube*[J]. BMC Genomi, 20(1): 464-478.